

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

THIS PAGE BLANK (USPTO)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Philippe MARLIERE, et al.

SERIAL NO.: NEW U.S. PCT APPLICATION

FILED: HEREWITH

INTERNATIONAL APPLICATION NO.: PCT/FR99/02628

INTERNATIONAL FILING DATE: 28 October 1999

FOR: PROCESS FOR PRODUCING CHEMICALLY DIVERSIFIED PROTEINS, IN VIVO, BY
INCORPORATING UNCONVENTIONAL AMINO ACIDS**REQUEST FOR PRIORITY UNDER 35 U.S.C. 119**
AND THE INTERNATIONAL CONVENTIONAssistant Commissioner for Patents
Washington, D.C. 20231

Sir:

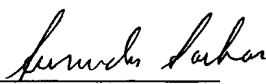
In the matter of the above-identified application for patent, notice is hereby given that the applicant claims as priority:

<u>COUNTRY</u>	<u>APPLICATION NO.</u>	<u>DAY/MONTH/YEAR</u>
FRANCE	98/13,533	28 October 1998

A certified copy of the corresponding Convention application(s) was submitted to the International Bureau in PCT Application No. **PCT/FR99/02628**. Receipt of the certified copy(s) by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.

Respectfully submitted,
OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.

**22850**


Norman F. Oblon
Attorney of Record
Registration No. 24,618
Surinder Sachar
Registration No. 34,423

(703) 413-3000
Fax No. (703) 413-2220
(OSMMN 1/97)

THIS PAGE BLANK (USPTO)

09/830669
PCT/FR 99/02628

REC'D 12 NOV 1999

WIPO PCT

BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION**COPIE OFFICIELLE****PRIORITY
DOCUMENT**SUBMITTED OR TRANSMITTED
BUT NOT IN COMPLIANCE WITH
RULE 17.1(a) OR (b)

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le **19 OCT. 1999**Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

Martine PLANCHE

**INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE****SIEGE**
26 bis, rue de Saint Petersburg
75800 PARIS Cédex 08
Téléphone : 01 53 04 53 04
Télécopie : 01 42 93 59 30

THIS PAGE BLANK (USPTO)



BREVET D'INVENTION, CERTIFICAT D'UTILITÉ

Code de la propriété intellectuelle-Livre VI

cerfa
N° 55-1328

REQUÊTE EN DÉLIVRANCE

Confirmation d'un dépôt par télécopie ☐

Cet imprimé est à remplir à l'encre noire en lettres capitales

26 bis, rue de Saint Pétersbourg
75800 Paris Cedex 08
Téléphone : 01 53 04 53 04 Télécopie : 01 42 93 59 30

Réserve à l'INPI

DATE DE REMISE DES PIÈCES

28.10.1998

N° D'ENREGISTREMENT NATIONAL

DÉPARTEMENT DE DÉPÔT

75

98 13533-

DATE DE DÉPÔT

28 OCT. 1998

1

NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE
À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE

CABINET REGIMBEAU
26, Avenue Kléber
75116 PARIS

n° du pouvoir permanent

références du correspondant

téléphone

237184 D17560 MIP

01 45 00 92 02

2 DEMANDE Nature du titre de propriété industrielle

☒ brevet d'invention

☐ demande divisionnaire

☐ certificat d'utilité

☐ transformation d'une demande
de brevet européen

☐ brevet d'invention

☐ certificat d'utilité n°

date

Établissement du rapport de recherche

☐ différé

☒ immédiat

Le demandeur, personne physique, requiert le paiement échelonné de la redevance

☐ oui

☐ non

Titre de l'invention (200 caractères maximum)

Procédé de production in vivo de protéines chimiquement diversifiées par incorporation
d'acides aminés non conventionnels

3 DEMANDEUR (S)

n° SIREN

code APE-NAF

Nom et prénoms (souligner le nom patronymique) ou dénomination

INSTITUT PASTEUR

Forme juridique

FONDATION RECONNUE D'UTILITE PUBLIQUE

Nationalité (s)

Française

Adresse (s) complète (s)

28, rue du Docteur Roux 75015 PARIS

Pays

FR

4 INVENTEUR (S) Les inventeurs sont les demandeurs

☐ oui

☒ non Si la réponse est non, fournir une désignation séparée

5 RÉDUCTION DU TAUX DES REDEVANCES

☐ requise pour la 1ère fois

☐ requise antérieurement au dépôt ; joindre copie de la décision d'admission

6 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE

pays d'origine

numéro

date de dépôt

nature de la demande

7 DIVISIONS

antérieures à la présente demande n°

date

n°

date

8 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE

(nom et qualité du signataire)

SIGNATURE DU PRÉPOSÉ À LA RÉCEPTION

SIGNATURE APRÈS ENREGISTREMENT DE LA DEMANDE À L'INPI

92-1001

DÉSIGNATION DE L'INVENTEUR

(si le demandeur n'est pas l'inventeur ou l'unique inventeur)

DEPARTEMENT DES BREVETS

26bis, rue de Saint-Petersbourg
75800 Paris Cédex 08
Tél. : 01 53 04 53 04 - Télécopie : 01 42 93 59 30

N° D'ENREGISTREMENT NATIONAL

9873533

TITRE DE L'INVENTION : Procédé de production in vivo de protéines chimiquement diversifiées par incorporation d'acides aminés non conventionnels

LE(S) SOUSSIGNÉ(S)

INSTITUT PASTEUR
28, rue du Docteur Roux 75015 PARIS

DÉSIGNE(NT) EN TANT QU'INVENTEUR(S) (indiquer nom, prénoms, adresse et souligner le nom patronymique) :

MARLIERE Philippe
Markgrafenstrasse 17
Konstanz, DE

DÖRING Volker
31, rue St Amand
75015 Paris, FR

MOOTZ Henning
Ziegelstrasse 11
35037 Marburg, DE

NOTA : A titre exceptionnel, le nom de l'inventeur peut être suivi de celui de la société à laquelle il appartient (société d'appartenance) lorsque celle-ci est différente de la société déposante ou titulaire.

Date et signature (s) du (des) demandeur (s) ou du mandataire

28 octobre 1998

CABINET REGIMBEAU

92-1001

DOCUMENT COMPORTANT DES MODIFICATIONS

PAGE(S) DE LA DESCRIPTION OU DES REVENDICATIONS OU PLANCHE(S) DE DESSIN			R.M.*	DATE DE LA CORRESPONDANCE	TAMPON DATEUR DU CORRECTEUR
Modifiée(s)	Supprimée(s)	Ajoutée(s)			
4				01.03.98	19 MAI 1999 - SR

Un changement apporté à la rédaction des revendications d'origine, sauf si celui-ci découle des dispositions de l'article R.612-36 du code de la Propriété Intellectuelle, est signalé par la mention «R.M.» (revendications modifiées).

La présente invention a pour objet des méthodes permettant à des cellules procaryotes ou eucaryotes d'acquérir la capacité de produire des protéines dont les séquences d'acides aminés comprennent au moins un acide aminé non conventionnel, des méthodes de sélection desdites
5 cellules, des procédés de production et de purification desdites protéines ainsi que les cellules et les protéines obtenues par les méthodes et procédés selon l'invention. L'invention comprend également les applications desdites
10 cellules et protéines dans différents domaines tels que le domaine thérapeutique, cosmétique, diagnostic ou de la biosynthèse ou la biodégradation de composé organique.

Un nombre croissant de protéines produites massivement par des organismes recombinants sont employées comme
15 catalyseurs dans l'industrie chimique ou comme agents thérapeutiques. La recherche de nouvelles protéines aux fonctions diversifiées est l'objet d'une activité intense, soit en criblant les protéines d'organismes extrêmophiles, soit en créant des variants protéiques par mutagenèse et
20 criblage. Toutefois, la variabilité chimique des protéines qui peuvent être produites dans des organismes vivants reste limitée par l'invariance du code génétique, c'est-à-dire restreinte aux combinaisons d'un jeu canonique de 20 acides aminés. Si la descendance des espèces naturelles
25 pouvait être progressivement remodelée dans le laboratoire de manière à adopter différents codes génétiques, l'évolution des protéines pourrait être redirigée et des sources artificielles de biodiversité ainsi établies.

La déviation expérimentale du code génétique est la
30 seule voie qui permettrait de surmonter cette limitation. Un autre code génétique pourrait spécifier un ensemble plus ou moins grand d'acides aminés, un ensemble substitué par des monomères non canoniques ou un ensemble d'acides aminés
canoniques parmi lesquels les codons sont redistribués. La
35 spécification d'acides aminés supplémentaires dans des lignées vivantes se prêterait à de multiples applications

dont la plus générique serait l'établissement d'une biodiversité artificielle.

L'incorporation, permanente ou transitoire, d'un seul acide aminé supplémentaire, portant un motif chimique qui
5 pourrait réagir sans modifier les acides aminés conventionnels, suffirait à fonder de nouvelles méthodes de fonctionnalisation de protéine. Ceci est justement l'objet de la présente invention.

L'invention a pour objet une méthode permettant à des
10 cellules d'acquérir la capacité de produire une protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel, caractérisée en ce qu'elle comprend les étapes suivantes :

- 15 a) la transformation desdites cellules par au moins une introduction d'une mutation faux-sens au niveau d'un codon cible d'un gène codant pour une protéine nécessaire à la croissance desdites cellules, ladite protéine synthétisée à partir du gène ainsi muté n'étant plus fonctionnelle ;
- 20 b) le cas échéant la culture des cellules obtenues à l'étape a) dans un milieu de culture contenant le nutriment exigé par la perte de fonctionnalité de ladite protéine ainsi mutée ; et
- 25 c) la culture des cellules obtenues à l'étape a) ou b) dans un milieu de culture contenant l'acide aminé codé par ledit codon cible.

On entendra désigner dans la présente description par le terme de protéine également les peptides ou les polypeptides, ainsi que les glycoprotéines correspondantes
30 lorsque lesdites protéines sont glycosylées.

On entendra également désigner dans la présente description par acide aminé non conventionnel tout acide aminé autre que les acides aminés incorporés par les
35 ribosomes au cours de la biosynthèse des protéines synthétisées par les organismes unicellulaires ou pluricellulaires procaryotes ou eucaryotes, ainsi que tout acide aminé incorporé à la place de l'acide aminé devant être

normalement incorporé à cette place au regard de la séquence nucléique traduite.

On entendra également désigner dans la présente description par mutation faux-sens, une mutation qui transforme un codon qui représente un acide aminé en un codon qui code pour un autre acide aminé, ce dernier, le cas échéant, ne pouvant remplacer l'acide aminé d'origine pour donner une protéine fonctionnelle dans la protéine à la place du résidu d'acide aminé d'origine.

On entendra également désigner dans la présente description par protéine nécessaire à la croissance de cellules, une protéine qui lorsqu'elle est synthétisée par les cellules de manière fonctionnelle permet auxdites cellules de croître dans des conditions de culture données et qui lorsqu'elle est synthétisée par les cellules de manière non fonctionnelle nécessite l'introduction d'un nutriment supplémentaire dans ledit milieu de culture donné pour permettre auxdites cellules de croître. De telles protéines non fonctionnelles peuvent être par exemple synthétisées par des cellules suite à des mutations conditionnelles telles qu'une mutation de type photosensible.

Afin d'illustrer par un exemple, mais sans s'y limiter, on peut citer notamment la protéine thymidylate synthase de *E. coli* qui présente un site catalytique occupé par la cystéine au niveau de la position 146 de sa séquence d'acides aminés et dont les mutations correspondantes du gène (*thyA*) causent une exigence nutritionnelle pour la thymine ou la thymidine, aucun autre acide aminé ne pouvant remplacer la cystéine à ce site.

On entendra désigner dans la présente description par codon cible, le codon de trois bases nucléotidiques transformé par la mutation faux-sens.

L'invention comprend également une méthode selon l'invention, caractérisée en ce que le milieu de culture de l'étape c) ne contient pas le nutriment exigé par la perte de fonctionnalité de ladite protéine mutée.

Selon l'invention, l'étape c) de culture desdites cellules peut comprendre une série de cultures desdites cellules dans un milieu de culture contenant l'acide aminé codé par ledit codon cible, chacune desdites cultures de la
5 série étant effectuée jusqu'à obtention de la phase stationnaire de croissance et suivie d'un lavage des cellules obtenues, le nombre de cultures de la série étant suffisant pour permettre la propagation de l'allèle correspondant audit gène muté.

10 L'invention concerne en outre une méthode selon l'invention, caractérisée en ce que la mutation faux-sens est choisie parmi les mutations faux-sens qui ne réversent spontanément qu'à une très faible fréquence, de l'ordre d'un organisme parmi au moins 10^{15} .

15 De préférence, la mutation faux-sens sera choisie parmi les mutations faux-sens qui transforment un codon cible d'un gène codant pour une protéine nécessaire à la croissance de ladite cellule en un codon qui comparative-ment au codon cible présente un changement d'au moins deux
20 bases, de manière plus préférée trois bases.

On préfère également les méthodes selon l'invention, caractérisées en ce que le codon cible code pour un acide aminé de faible volume stérique et/ou amphiphile et/ou de volume stérique inférieur ou sensiblement égal au volume
25 stérique de l'acide aminé codé par la mutation faux-sens.

Parmi les codons cibles, on préfère notamment les codons cibles codant pour la cystéine et les mutations faux-sens choisies parmi les mutations faux-sens qui transforment un codon cible en un codon codant pour la
30 valine ou l'isoleucine.

L'invention concerne en outre une méthode selon l'invention, caractérisée en ce que l'étape a) de
~~transformation desdites cellules est réalisée au moyen d'un~~
vecteur comprenant une séquence dudit gène codant pour une
35 protéine nécessaire à la croissance desdites cellules comportant ladite mutation faux-sens, notamment au moyen d'un vecteur plasmidique.

De tels vecteurs seront préparés selon les méthodes couramment utilisées par l'homme du métier, et les clones en résultant peuvent être introduits dans lesdites cellules par des méthodes usuelles de recombinaison génétique telles que par exemple la lipofection, l'électroporation ou le choc thermique.

Sous un autre aspect, l'invention a pour objet une méthode de sélection de cellules capables de produire une protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel caractérisée en ce qu'elle comprend les étapes a), le cas échéant b), et c) d'une méthode selon l'invention, et la sélection des cellules capables de croître à l'étape c).

De manière préférée, la méthode de sélection de cellules selon l'invention, comprendra en outre une étape d) de culture des cellules obtenues à l'étape c) dans un milieu de culture contenant ledit acide aminé codé par ledit codon cible, la concentration dudit acide aminé pouvant être à une concentration supérieure à la concentration dudit acide aminé utilisée à l'étape c), et le choix des cellules sensibles à la concentration dudit acide aminé utilisée à l'étape d).

On entend désigner par cellule sensible à un composé chimique ou biochimique ou à une concentration donnée dudit composé, une cellule dont la croissance est partiellement ou totalement inhibée lorsqu'elle est cultivée dans un milieu de culture contenant ledit composé chimique ou biochimique ou ladite concentration dudit composé.

L'invention comprend également une méthode de sélection de cellules selon l'invention, caractérisée en ce que l'aminoacyl-tRNA synthétase reconnaissant l'acide aminé codé par ladite mutation faux-sens desdites cellules sélectionnées est capable de charger sur un de ses tRNA associés un acide aminé non conventionnel ou un acide aminé autre que ledit acide aminé codé par ladite mutation faux-sens.

On entendra désigner dans la présente description par tRNA associé, un tRNA qui est reconnu par l'aminoacyl-tRNA

synthétase reconnaissant un acide aminé et qui peut transférer ledit acide aminé.

L'invention comprend en outre une méthode de sélection de cellules mutantes selon l'invention, caractérisée en ce
5 que la séquence nucléique du gène codant pour ladite aminoacyl-tRNA synthétase comporte au moins une mutation comparée à la séquence du gène sauvage correspondant, ladite mutation n'ayant pas été introduite par une technique de recombinaison génétique.

10 Selon un autre aspect, l'invention a pour objet les cellules procaryotes ou eucaryotes obtenues par une méthode selon l'invention.

Parmi les cellules utilisables à ces fins, on peut citer bien entendu les cellules bactériennes telles que E.
15 coli, mais également les cellules de levure, de même que les cellules animales, en particulier les cultures de cellules de mammifères, comme notamment les cellules d'ovaire de hamster chinois (CHO), mais également les cellules d'insectes.

20 L'invention concerne également les cellules procaryotes ou eucaryotes isolées capables de produire une protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel, caractérisées en ce qu'elles comprennent une aminoacyl-tRNA synthétase
25 reconnaissant un acide aminé donné capable de charger sur un de ses tRNA associés un acide aminé non conventionnel ou un acide aminé autre que ledit acide aminé donné, et en ce que la séquence nucléique du gène codant pour ladite aminoacyl-tRNA synthétase comporte au moins une mutation
30 comparée à la séquence du gène sauvage correspondant, ladite mutation n'ayant pas été introduite par une technique de recombinaison génétique.

Ainsi, l'invention concerne une méthode de sélection de cellules basée sur la constitution par la cellule d'une
35 voie métabolique nécessaire à sa croissance permettant d'obtenir des cellules capables de produire un acyl-tRNA

non canonique capable de charger un acide aminé non conventionnel.

Parmi les cellules selon l'invention, on préfère les cellules de bactérie, caractérisées en ce qu'elles sont choisies parmi les cellules suivantes déposées à la CNCM (Collection Nationale de Culture de Microorganismes, Paris, France) :

- a) souche *E. coli* déposée à la CNCM sous le N° I-2025 le 25 mai 1998,
- 10 b) souche *E. coli* déposée à la CNCM sous le N° I-2026 le 25 mai 1998, et
- c) souche *E. coli* déposée à la CNCM sous le N° I-2027 le 25 mai 1998.

15 L'invention comprend en outre l'utilisation d'une méthode ou d'une cellule selon l'invention pour la production de protéine, notamment recombinante, dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel.

20 Sous un autre aspect, l'invention concerne un procédé de production d'une protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel, caractérisé en ce qu'il comprend les étapes suivantes :

- a) le cas échéant, la sélection d'une cellule par une méthode selon l'invention ;
- 25 b) la culture de ladite cellule sélectionnée à l'étape a) ou d'une cellule selon l'invention dans un milieu de culture et des conditions de culture permettant la croissance de ladite cellule ; et
- 30 c) l'isolement de ladite protéine comprenant au moins un acide aminé non conventionnel à partir du surnageant de culture et/ou du culot cellulaire obtenu à l'étape b).

Parmi les protéines pouvant être produites par un procédé selon l'invention, on peut mentionner, mais sans s'y limiter, les protéines qui par l'incorporation d'au moins un acide aminé non conventionnel permettent d'obtenir une activité recherchée qu'une protéine dont la séquence

comporte uniquement des acides aminés conventionnels ne permet pas d'obtenir. Par activité, on entend désigner de manière générale toute activité telle qu'une activité physiologique ou biologique relative aux organismes uni- ou pluricellulaires, même partielle, comme par exemple une activité structurale ou biochimique, par exemple enzymatique, antigénique, de type anticorps ou de modulation, de régulation ou d'inhibition d'activité biologique, ou encore telle qu'elle permette sa mise en oeuvre dans un procédé de biosynthèse ou de biodégradation de composés chimiques ou biochimiques.

On peut également mentionner parmi les protéines pouvant être produites par un procédé selon l'invention, les protéines dont l'incorporation d'au moins un acide aminé non conventionnel est effectuée de telle manière qu'il n'en résulte pas de modification approfondie de l'activité biologique de la protéine non modifiée correspondante. Outre l'activité biologique conservée de la protéine non modifiée correspondante, ces protéines selon l'invention présenteront un acide aminé non conventionnel dont les propriétés spécifiques pourront être avantageusement exploitées.

Parmi les propriétés spécifiques conférées par la présence d'un acide aminé non conventionnel, on peut citer en particulier les propriétés liées à la présence d'un groupe fonctionnel sur ledit acide aminé non conventionnel capable de réagir facilement et de manière spécifique avec un composé chimique ou biochimique dans des conditions permettant de ne pas altérer l'activité de la protéine ou évitant la modification des acides aminés conventionnels.

La présence de ce groupe fonctionnel spécifique pourra avantageusement être utilisée par exemple pour :

-
- (i) purifier toute protéine, notamment recombinante, incorporant ledit acide aminé non conventionnel ;
 - (ii) coupler une telle protéine à un support solide ;

(iii) coupler à une telle protéine des molécules capables d'être détectées, telles que des sondes spectroscopiques de nature variée ;

5 (iv) coupler à une telle protéine des polymères lipophiles ou hydrophiles permettant de les solubiliser dans des solvants ou de faire écran à la reconnaissance par des anticorps ;

(v) coupler une telle protéine à un polynucléotide ;

10 (vi) coupler une telle protéine à un composé chimique ou biochimique dont la présence permet d'augmenter, de diminuer, de moduler, de réguler ou de cibler l'activité biologique de ladite protéine, ou de modifier sa biodisponibilité en tant que composé à usage thérapeutique ; ou encore

15 (vii) fixer de manière permanente à une telle protéine un coenzyme qui sinon diffuserait en solution.

Selon la présente invention, l'incorporation d'au moins un acide aminé non conventionnel pourra porter sur des acides aminés à l'origine d'une spécificité ou de
20 l'activité, ou à l'origine de la conformation structurale, de la charge, de l'hydrophobicité ou de la capacité de multimérisation de la protéine non modifiée correspondante. Ainsi, pourront être créées des protéines d'activité équivalente, augmentée ou diminuée, ou de spécificité
25 équivalente, plus étroite, ou plus large que la protéine à acides aminés conventionnels non modifiée correspondante.

Par protéine non modifiée, on entend désigner la protéine sauvage ou recombinante constituée d'acides aminés conventionnels et dont est issue la protéine comprenant
30 l'acide aminé non conventionnel.

De préférence, le procédé de production selon l'invention est caractérisé en ce que ledit milieu de culture de l'étape b) permettant la croissance de ladite cellule contient ledit acide aminé non conventionnel ou un
35 de ses précurseurs.

Selon un mode particulier, un procédé de production selon l'invention est caractérisé en ce que ledit acide

aminé non conventionnel est synthétisé par ladite cellule, la synthèse dudit acide aminé non conventionnel pouvant être augmentée par modification génétique de ladite cellule.

5 L'invention concerne en outre un procédé de production d'une protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel selon l'invention, caractérisé en ce que ladite cellule est auxotrophe pour l'acide aminé codé par ledit codon cible.

10 Sont également compris dans la présente invention, les procédés selon l'invention, caractérisés en ce que ladite cellule comprend un gène d'intérêt homologue ou hétérologue dont la séquence codante comporte au moins un codon cible.

D'une manière générale, le gène d'intérêt codera pour
15 un ARN messager qui sera ensuite traduit en protéine d'intérêt.

Le gène d'intérêt peut être isolé par toute technique conventionnelle telle que clonage, PCR (Polymerase Chain Reaction) ou encore synthétisé chimiquement. Il peut être
20 de type génomique (muni d'un ou plusieurs introns) ou ADN complémentaire (ADNc). La protéine d'intérêt peut être constituée par une protéine mature, un précurseur et, notamment un précurseur destiné à être sécrété et comprenant un peptide signal, une protéine tronquée, une
25 protéine chimère provenant de la fusion de séquences d'origines diverses ou encore une protéine mutée présentant des propriétés biologiques améliorées et/ou modifiées.

De manière générale, le gène d'intérêt homologue ou hétérologue pourra être choisi parmi les gènes codant pour
30 toute protéine utilisable comme composé thérapeutique ou cosmétique, ou comme réactif de diagnostic, ou encore comme composé pouvant être mis en oeuvre dans un procédé de biosynthèse ou de biodégradation.

A titre d'exemples, on peut citer les gènes d'intérêt
35 codant pour les protéines d'intérêt suivantes :

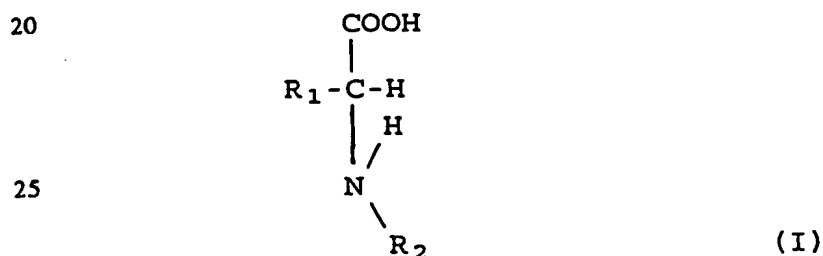
- cytokines ou lymphokines (interférons α , β et γ , interleukines et notamment l'IL-2, l'IL-6, l'IL-10 ou l'IL-12, facteurs nécrosant des tumeurs (TNF), facteurs stimulateurs de colonies (GM-CSF, C-CSF, M-CSF, ...) ;
 - 5 - récepteurs cellulaires ou nucléaires, notamment ceux reconnus par des organismes pathogènes (virus, bactéries, ou parasites) ou leurs ligands ;
 - protéines impliquées dans une maladie génétique (facteur VII, facteur VIII, facteur IX, dystrophine ou minidystrophine, insuline, protéine CFTR (Cystic Fibrosis Transmembrane Conductance Regulator), hormones de
 - 10 croissance (hGH) ;
 - enzymes (uréase, rénine, thrombine, ...) ou toutes enzymes impliquées dans le métabolisme ou la biosynthèse des
 - 15 protéines, des lipides, des acides nucléiques, des sucres, des acides aminés, des acides gras ou des nucléotides ;
 - inhibiteurs d'enzymes (α 1-antitrypsine, antithrombine III, inhibiteurs de protéases virales, ...) ;
 - 20 - composés à effet anti-tumoral capables d'inhiber au moins partiellement l'initiation ou la progression de tumeurs ou cancers (anticorps, inhibiteurs agissant au niveau de la division cellulaire ou des signaux de transduction, produits d'expression des gènes suppresseurs de tumeurs,
 - 25 par exemple p53 ou Rb, protéines stimulant le système immunitaire, ...) ;
 - protéines du complexe majeur d'histocompatibilité des classes I ou II ou protéines régulatrices agissant sur l'expression des gènes correspondants ;
 - 30 - protéines capables d'inhiber une infection virale, bactérienne ou parasitaire ou son développement
-
- (protéines antigéniques ayant des propriétés immunogènes, épitopes antigéniques, anticorps, ...) ;
- toxines telles que la ricine, toxine cholérique,
 - 35 diphtérique, ... ou immunotoxines ;

- marqueurs (β -galactosidase, peroxydase, ...) ; et
- luciférase, GFP (green fluorescent protein), etc..

L'invention comprend en outre un procédé de production d'une protéine selon l'invention, caractérisé en ce que le milieu de culture de l'étape b) comprend en outre les composés nécessaires à l'induction de la synthèse de la protéine codée par ledit gène d'intérêt. Ces composés sont connus de l'homme du métier et dépendent notamment de la cellule et du gène homologue ou hétérologue sélectionnés.

L'invention concerne également un procédé selon l'invention, caractérisé en ce que l'activité biologique de la protéine codée par ledit gène d'intérêt est au moins partiellement conservée après l'incorporation dudit acide aminé non conventionnel au niveau du codon cible dudit gène d'intérêt.

L'invention concerne en outre un procédé selon l'invention, caractérisé en ce que l'acide aminé non conventionnel est choisi parmi les acides aminés non conventionnels de formule I et de configuration L



dans laquelle :

R_1 ou R_2 représente des radicaux contenant un groupe fonctionnel capable de réagir de manière sélective, de préférence choisi parmi les groupes aldéhyde, cétone, éthényle, éthynyle ou nitrile.

Parmi ces groupes, on préfère particulièrement le groupe oxo (aldéhyde ou cétone) à réactivité sélective qui faciliterait la fonctionnalisation chimique des protéines.

D'autres groupes simples comme le groupe éthynyle se prêteraient également à des réactions sélectives. Un vaste corpus d'expériences conduites à l'aide de systèmes de traduction acellulaire (ex vivo) et d'acyl-tRNAs

synthétisés in vitro, a démontré qu'une grande variété de groupes acyles pouvaient être transférés sur le ribosome en réponse à un codon lu par le tRNA. En bref, les modifications latérales des acides aminés semblent toutes
 5 compatibles avec la traduction (il n'a à ce jour pas été trouvé d'acide aminé dont la chaîne latérale serait assez encombrante pour bloquer la traduction) ; des substitutions du motif amino en alkyl-amino, en hydroxy et en hydrazino sont compatibles avec la chimie de transpeptidation
 10 catalysée par le ribosome (Bain et al. 1991) (il est connu que le ribosome peut former des polyesters en plus des polyamides conventionnels) ; la substitution de l'hydrogène alpha du motif $H_2NCH(R)-COOH$ par un groupe alkyle (méthyle) ou l'inversion de configuration au carbone alpha (D-amino-
 15 acides) ne sont par contre pas acceptées par le ribosome.

L'invention a également pour objet un procédé selon l'invention, pour la fonctionnalisation de protéine.

L'invention concerne également un procédé de purification de protéine, caractérisé en ce qu'il comprend
 20 les étapes suivantes :

- a) l'incorporation dans la séquence d'acides aminés de ladite protéine d'un acide aminé non conventionnel contenant un groupe fonctionnel capable de réagir de manière sélective par un procédé selon l'invention ;
- 25 b) la mise en contact de la solution contenant la protéine obtenue à l'étape a) avec un support comprenant un composé capable de réagir spécifiquement avec ledit groupe fonctionnel et de fixer spécifiquement ladite protéine; et
- 30 c) l'isolement de ladite protéine fixée sur le support.

Les procédés de purification de protéine, naturelle ou recombinante, utilisés habituellement par l'homme du métier font appel généralement à des méthodes utilisées
 35 individuellement ou en combinaison, telles que le fractionnement, les méthodes de chromatographie, les techniques d'immunoaffinité à l'aide d'anticorps mono ou polyclonaux spécifiques, etc.. Ces méthodes sont parfois longues et

fastidieuses et ne permettent pas toujours d'obtenir l'activité spécifique, le taux et le rendement de purification voulus. La présence d'un groupe fonctionnel spécifique sur la protéine à purifier capable de réagir
5 sélectivement avec le support de purification sans altérer l'activité de la protéine faciliterait grandement la purification de protéine nécessaire à leur utilisation.

L'invention concerne également un procédé de fixation d'une protéine sur un composé chimique ou biochimique,
10 caractérisé en ce qu'il comprend les étapes suivantes :

- a) l'incorporation dans la séquence d'acides aminés de ladite protéine par un procédé selon l'invention d'un acide aminé non conventionnel contenant un groupe fonctionnel capable de réagir de manière sélective ;
- 15 b) la mise en contact de la protéine obtenue à l'étape a) avec ledit composé chimique ou biochimique comprenant un groupe capable de réagir spécifiquement avec ledit groupe fonctionnel dans un milieu permettant la réaction.

20 De préférence, la fixation d'une protéine sur un composé chimique ou biochimique est une fixation par liaison covalente.

Les composés chimiques ou biochimiques pouvant être utilisés dans ledit procédé de fixation selon l'invention
25 pourront être choisis parmi tous les composés capables de réagir avec le groupe fonctionnel de l'acide aminé non conventionnel incorporé.

On entend désigner dans la présente description par complexe protéique le produit obtenu à l'étape b) du
30 procédé décrit ci-dessus comprenant une protéine selon l'invention fixée sur un composé chimique ou biochimique.

L'invention comprend aussi un procédé selon
~~l'invention caractérisé en ce que ledit composé chimique ou~~
biochimique est lui-même fixé sur un support solide ou est
35 un composé constitutif d'un support solide.

L'invention concerne en outre un procédé selon l'invention pour la préparation d'un complexe protéique.

De préférence, l'invention comprend les procédés de l'invention, caractérisés en ce que ladite protéine fixée ou ledit composé chimique ou biochimique est choisi parmi les composés thérapeutiques, cosmétiques ou diagnostiques.

5 Ladite protéine fixée sera choisie en particulier parmi les protéines dont la séquence d'acides aminés comprend un acide aminé non conventionnel selon un procédé de l'invention, et dont la protéine non modifiée correspondante sauvage ou recombinante est choisie parmi
10 les protéines utilisables comme composés thérapeutiques, cosmétiques ou comme réactifs de diagnostic.

De préférence, les procédés selon l'invention, sont caractérisés en ce que le composé chimique ou biochimique est choisi parmi les composés capables de modifier
15 l'activité biologique de la protéine fixée.

On entend désigner par composés capables de modifier l'activité biologique d'un autre composé, un composé capable d'augmenter, de diminuer, de réguler l'activité biologique dudit autre composé.

20 L'invention comprend également un procédé selon l'invention, caractérisé en ce que le composé chimique ou biochimique est choisi parmi les composés dont l'activité biologique peut être modifiée par la protéine fixée.

L'invention comprend également un procédé selon
25 l'invention, caractérisé en ce que le composé chimique ou biochimique est choisi parmi les composés comprenant une protéine, un polynucléotide, un acide gras, un sucre ou un polymère naturel ou synthétique.

Selon un autre aspect, l'invention a pour objet les
30 protéines, en particulier recombinantes, et les complexes protéiques obtenus par un procédé selon l'invention.

L'invention comprend également une méthode de
~~sélection de composés capables de se lier à une protéine~~
selon l'invention ou capables de se lier au composé
35 chimique ou biochimique du complexe protéique selon l'invention. Parmi ces méthodes, on peut citer comme

exemple une méthode caractérisée en ce qu'elle comprend les étapes suivantes :

a) la mise en contact dudit composé susceptible d'être sélectionné avec la protéine ou le complexe protéique selon l'invention, ladite protéine ou complexe protéique pouvant être notamment fixée sur un support solide ;

b) la détermination de la capacité dudit composé à se lier avec la protéine ou le complexe protéique selon l'invention.

Les composés susceptibles d'être sélectionnés peuvent être des composés organiques tels que des protéines ou hydrates de carbone ou tous autres composés organiques ou inorganiques déjà connus, ou des composés organiques nouveaux élaborés à partir de techniques de modélisation moléculaire et obtenus par synthèse chimique ou biochimique, ces techniques étant connues de l'homme de l'art.

Les cellules selon l'invention pourront également avantageusement servir de modèle et être utilisés dans des procédés pour étudier, identifier et/ou sélectionner des protéines selon l'invention ou des composés susceptibles de posséder une activité recherchée.

L'invention concerne également l'utilisation d'une protéine ou d'un complexe protéique selon l'invention comme réactif de diagnostic ainsi que les procédés de diagnostic, notamment pour la détection, l'identification, la localisation et/ou le dosage spécifique de polypeptide ou de polynucléotide, mettant en oeuvre une protéine ou un complexe protéique selon l'invention.

Sont en effet comprises dans les protéines selon l'invention, des protéines ayant incorporé au moins un acide aminé non conventionnel et ayant conservé partiellement ou totalement l'activité initiale des protéines sauvages ou recombinantes non modifiées correspondantes, telles que des anticorps, des antigènes,

des enzymes ou leurs fragments biologiquement actifs connus pour être utilisés dans des procédés de diagnostic.

De la même manière, sont compris dans les complexes protéiques selon l'invention, des complexes protéiques
5 formés à partir d'une protéine selon l'invention et un composé chimique ou biochimique tels que des complexes comprenant un anticorps, un antigène ou une sonde oligonucléotidique couplé à une enzyme, à un substrat ou à une molécule capable d'être détectée.

10 Parmi les procédés de diagnostic selon l'invention, on peut citer par exemple les procédés comprenant les étapes suivantes :

a) la mise en contact de l'échantillon biologique susceptible de contenir le composé recherché avec une
15 protéine ou un complexe protéique selon l'invention, ladite protéine ou complexe protéique pouvant être notamment fixée sur un support solide ; et

b) la mise en évidence, l'identification, la localisation et/ou le dosage du complexe formé entre le
20 composé recherché et une protéine ou un complexe protéique selon l'invention.

L'homme du métier saura adapter avec les protéines ou les complexes protéiques selon l'invention les procédés de diagnostic standards connus.

25 Les techniques et les réactifs spécifiques permettant la mise en évidence, l'identification, la localisation et/ou le dosage du complexe formé que l'on pourra utiliser dans les procédés de l'invention, sont bien connus également de l'homme du métier et sont, par exemple, les
30 techniques ELISA, RIA, d'immunofluorescence, de PCR ou d'autres techniques d'amplification d'un acide nucléique cible connus de l'homme de l'art.

L'invention concerne également un kit ou nécessaire de diagnostic, notamment pour la détection, l'identification,
35 la localisation et/ou le dosage spécifique de protéine ou de polynucléotide caractérisé en ce qu'il contient une protéine ou un complexe protéique selon l'invention.

L'invention concerne en outre l'utilisation d'une protéine, d'un complexe protéique ou d'une cellule selon l'invention pour la préparation d'une composition pharmaceutique ou cosmétique.

5 L'invention a enfin pour objet une composition pharmaceutique ou cosmétique comprenant une protéine, un complexe protéique ou une cellule selon l'invention.

D'autres caractéristiques et avantages de l'invention apparaissent dans la suite de la description avec les
10 exemples ci-après.

EXEMPLES

EXEMPLE 1 : Construction d'une souche d'E. coli
15 comportant une mutation faux-sens Cys->Val au site actif de la thymidylate synthase et créant une exigence nutritionnelle pour la thymine, la thymidine ou la cystéine.

Les allèles artificiels du gène thyA sont construits
20 par mutagenèse dirigée du plasmide pTS0 (Lemeignan et al 1993), qui dérive du plasmide pTZ18R (BioRad) par insertion du gène thyA sauvage d'E. coli. La mutagenèse dirigée à l'aide d'un oligonucleotide est réalisée selon la méthode décrite par Kunkel et coll. (1987) sur le phagémide pTS0.
25 La préparation de la matrice simple brin de pTS0, amplifiée dans la souche RZ1032 (Kunkel et coll., 1987) (Hfr KL16 P045 [lysA(61-62)] dut1 ung1 thil relA1 supE44 zbd-279::Tn10) est réalisée selon le protocole décrit par Sambrook et coll. (1989). Un oligonucléotide phosphorylé en 5' (acheté
30 à la société Genome Express) est utilisé comme amorce mutagène :

Oligodéoxynucléotide 1 :

5'-pTGGATAAAATGGCGCTGGCACCGGTACATGCATTCTTCCAGTTCTATGT

L'hybridation de cet oligonucléotide avec la matrice
35 simple brin dans chacune des deux constructions est réalisée avec 10 ng d'oligonucléotide et 0,2 µl de matrice dans un volume de 10 µl d'une solution tampon contenant

20 mM de Tris-HCl pH 7,5, 2 mM d'EDTA et 50 mM de chlorure de sodium. Les tubes sont incubés 5 mn à 70°C puis refroidis progressivement jusqu'à 30°C. A ce mélange est alors ajouté 0,5 mM de chacun des dNTP, 1 mM d'ATP, 10 mM de Tris-HCl à pH 7,5, 10 mM de chlorure de magnésium, 2 mM de dithiothréitol et 1 unité de chacune des deux enzymes du phage T4 ADN ligase et ADN polymérase. Ce mélange réactionnel d'un volume final de 20 µl est incubé 60 min à 37°C dont 5 µl sont ensuite utilisés pour transformer des cellules compétentes de la souche GT869 (Parsot, C. 1986) (thrB1004 pro thi strA hsdS lacZ ΔM15 [F' lacZ ΔM15 lacIq traD36 proA+ proB+]) d'E.coli K12 suivant la méthode décrite par Sambrook et coll. (1989). Les cellules transformées sont étalées sur des boîtes de Pétri contenant du milieu LB additionné de 100 mg/l de carbénicilline. Douze clones résistants à l'antibiotique sont réisolés sur le même milieu. L'ADN simple-brin correspondant aux phagémides de ces clones est préparé et séquencé selon la méthode didéoxy (Sanger et coll., 1977). Le kit de séquençage M13 (Boehringer Mannheim, Mannheim, Allemagne) et la désoxyadénosine 5'-(α-thio)triphosphate (1300 Ci/mmol, Amersham) sont combinées selon les indications des fournisseurs. Quatre amorces sont utilisées pour déterminer la séquence des allèles de thyA :

- 25 Oligodéoxynucléotide 3 : 5'GGTGTGATCATGATGGTC
 Oligodéoxynucléotide 4 : 5'CCTGCAAGATGGATTCCC
 Oligodéoxynucléotide 5 : 5'CGCGCCGCATTATTGTTTC
 Oligodéoxynucléotide 6 : 5'GTCTGGACCGGTGGCGACA

Le plasmide pTS1 ainsi obtenu propage l'allèle thyA:Val146, dans lequel la position 146 occupée dans le gène thyA sauvage par le codon UGC de la cystéine est occupée par le codon GUA de la valine. Le plasmide pTS1 est introduit par transformation, réalisée selon la méthode de Sambrook et coll. (1989), dans la souche ΔthyA d'E. coli K12, β1308 (Lemeignan et coll., 1993), dont le gène chromosomique de la thymidylate synthase, thyA, est délété.

La souche transformée portant l'allèle plasmidique thyA:Val146, β 5366, se montre incapable de croître sans que de la thymine ou de la thymidine soit ajoutée au milieu de culture, tout comme la souche β 1308 dont elle dérive. Par

5 contre, la souche β 5366 montre une croissance marginale à 30°C sur gradient de diffusion de cystéine, réalisé dans des boîtes de Pétri contenant 25 ml de milieu minéral MS glucose à partir d'un puits central comportant 0,1 ml d'une solution 400 mM de L-cystéine. Dans les mêmes conditions,

10 la souche β 1308 ne donne lieu à aucune croissance détectable. Ainsi, la mutation faux-sens convertissant la cystéine catalytique à la position 146 de la thymidylate synthase en valine peut être partiellement supprimée par apport massif de cystéine exogène. L'addition de 0,1 mM de

15 valine au milieu des boîtes de Pétri abolit la croissance de la souche β 5366 sur gradient de cystéine. Ainsi, tout se passe comme si la cystéine pouvait s'infiltrer dans le site actif de la valyl-tRNA synthétase pour former des Cys-tRNA^{Val} erronés, aptes à corriger la substitution de la

20 cystéine par la valine dans le site actif de la thymidylate synthase. L'excès de valine préviendrait la formation de ces Cys-tRNA^{Val} erronés.

Exemple 2 : Construction d'une souche d'E. coli

25 comportant une mutation faux-sens Cys->Ile au site actif de la thymidylate synthase et créant une exigence nutritionnelle pour la thymine, la thymidine ou la cystéine.

La construction correspondante est également conduite

30 pour remplacer la cystéine en position 146 de la thymidylate synthase, par mutagénèse dirigée à l'aide de l'oligonucléotide 2, suivant le même protocole que dans l'Exemple 1.

Oligodéoxynucléotide 2 :

35 5'pTGGATAAAATGGCGCTGGCACCGATACATGCATTCTTCCAGTTCTATGT

Le plasmide pTS2 ainsi obtenu propage l'allèle thyA:Ile146, dans lequel la position 146 occupée dans le gène thyA sauvage par le codon UGC de la cystéine est occupée par le codon AUA de l'isoleucine. La souche
 5 propageant l'allèle plasmidique thyA:Ile146, β 5274, se révèle exiger l'apport nutritionnel de thymine, de thymidine ou de cystéine en excès, tout comme la souche β 5366. La suppression phénotypique de la souche β 5274 par la cystéine est abolie par 0,1 mM d'isoleucine, tout comme
 10 celle de la souche β 5366 par la valine. Ainsi, tout se passe comme si l'isoleucyl-tRNA synthétase était capable de former des Cys-tRNA^{Ile} erronés en présence d'un excès de cystéine, et que cette formation erronée était prévenue par la présence d'un excès d'isoleucine.

15

Exemple 3 : Sélection de mutants du code génétique mésincorporant la cystéine au lieu de la valine par culture sérielle en liquide et caractérisation génétique des mutants de la valyl-tRNA synthétase ainsi obtenus.

20

La souche β 5366 portant l'allèle faux-sens thyA:Val146 sur le plasmide pTS1 est cultivée en milieu minéral MS glucose (2 g/l, Richaud et coll., 1993) supplémenté avec 0,3 mM de thymidine pendant 24 h à 30°C en aérobiose. Les
 25 cellules sont ensuite lavées deux fois avec du milieu minéral MS désoxygéné. Un milieu nutritif désoxygéné contenant 10 ml de milieu minéral MS glucose additionné de 1,5 mM de cystéine est inoculé au 1/100 à l'aide des cellules lavées. Les cellules sont ensuite cultivées en
 30 anaérobiose pendant 24 h à 30°C et un tube frais contenant 10 ml de milieu minéral MS glucose cystéine désoxygéné est inoculé avec une dilution au 1/100 de la culture en phase stationnaire précédente. Cette procédure est répétée 26
 fois. A l'issue de cette propagation sérielle, 12 clones de
 35 la culture liquide sont isolés sur boîtes de milieu minéral MS glucose (2 g/l) thymidine (0,3 mM) en aérobiose et

sauvegardés en suspension dans le même milieu liquide à - 80°C. Les douze clones sont testés sur des boîtes contenant du milieu minéral MS glucose additionné de facteurs nutritionnels. Tous ces clones se révèlent exiger la thymine ou la thymidine comme facteur de croissance, à moins que de la cystéine soit présente dans le milieu de culture, à une concentration d'au moins 1,5 mM.

Deux tels clones sont choisis pour une caractérisation génétique approfondie, β 8144 et β 8146. Des expériences de transduction par le phage P1 du caractère de résistance à la kanamycine, introduit dans le locus *nrdD* voisin du gène *valS* de la valyl-tRNA synthétase (97 mn du chromosome d'*E. coli* K12) sont effectuées à l'aide des souches β 8144 et β 8146. Dans les deux cas, environ la moitié des transductants résistant à la kanamycine montrent également une dépendance nutritionnelle pour la thymidine suppressible par la cystéine exogène à la concentration d'au moins 1,5 mM. Cette proportion est en accord avec la distance génétique entre les gènes *valS* et *nrdD* (0,4 mn) et laisse supposer que le phénotype de suppression de la mutation faux-sens Cys->Val au site actif de la thymidylate synthase par de faibles concentrations de cystéine est causé par l'altération génétique du locus *valS*.

La fixation d'une altération génétique dans le gène *valS* des souches adaptées est confirmée par séquençage de ce locus : un A changé en C cause le remplacement de la lysine à la position 277 par la glutamine dans les deux souches adaptées β 8144 et β 8146. Le séquençage est effectué sur une matrice obtenue par réaction de polymérisation en chaîne (PCR) réalisée dans des conditions décrites par Sambrook et coll. (1989). La réaction d'amplification est réalisée dans 100 μ l d'une solution contenant 10 ng d'ADN génomique des souches β 8144 ou β 8146, 20 pmoles de chaque amorce, 40 nmoles d'un mélange équimolaire des 4 désoxynucléotides triphosphate, 10 μ l d'un tampon composé de 100 mM Tris-HCl pH 8,3, 500 mM KCl et 20 mM MgCl₂, en

présence de 1 à 2 unités de Vent polymérase (Biolabs). Pour chaque réaction, 30 cycles de polymérisation sont accomplis, en utilisant un amplificateur d'ADN (Perkin-Elmer Cetus), comme suit : la dénaturation est effectuée à 94°C pendant 5 min pour le 1er cycle et 1 min pour les cycles suivants, l'hybridation à 58°C pendant 1 min et l'élongation à 72°C pendant 3 min pour les 29 premiers cycles et pendant 10 min pour le dernier cycle. Les oligonucléotides 7 et 8 sont utilisés pour l'amplification du gène.

Oligodéoxynucléotide 7 :

5'GGGGAATTTCGGTGTGTGAAATTGCCGCAGAACG

Oligodéoxynucléotide 8 :

5'GGCAAGCTTCCAGTATTTACAGGGGAGTTATGC

Les fragments de PCR ainsi obtenus sont purifiés en utilisant le kit QIAquick (Qiagen) et transmis à la société Genaxis pour en déterminer la séquence.

Exemple 4 : Suppression phénotypique par des précurseurs métabolique de la cystéine.

L'exigence nutritionnelle en cystéine des souches adaptées β 8144 et β 8146 est mise à profit pour caractériser des précurseurs métaboliques qui puissent se substituer à la cystéine dans le milieu de culture sans donner lieu à la dégradation par oxydation. La S-carbamyl-L-cystéine (3 mM), la S-méthyl-L-cystéine (3 mM) et l'acide L-thiazolidine-4-carboxylate (2 mM) se révèlent capables de remplacer la cystéine comme facteur de croissance des souches adaptées β 8144 et β 8146, au lieu de la thymidine ou de la thymine. Les mêmes composés se révèlent capables de satisfaire l'exigence en cystéine d'un mutant *cysN::kan* (souche JT1, procurée par M. Berlyn, Coli Genetic Stock Center, Yale University, USA (Levh et al., 1988)). Toutefois, l'addition d'aucune de ces substances ne permet la croissance de la souche β 1308 portant une délétion chromosomique du gène

thyA de la thymidylate synthase, excluant ainsi leur contamination par des traces de thymine ou de thymidine.

Exemple 5 : Sélection de mutants du code génétique
 5 mésincorporant la cystéine au lieu de la valine par isolement sur milieu solide et caractérisation génétique des mutants de la valyl-tRNA synthétase mésincorporant la cystéine.

10 La souche β 5366 portant l'allèle faux-sens thyA:Val146 sur le plasmide pTS1 est transduite avec un lysat du phage P1 récolté sur une souche auxiliaire d'*E. coli* (β 7170, Bouzon et al., 1997) dans le chromosome de laquelle un marqueur de résistance à la kanamycine avait été introduit
 15 au locus nrdD, voisin du locus valS du gène de la valyl-tRNA synthétase, produisant ainsi la souche β 5419. Un allèle mutateur du gène dnaQ est introduit extemporanément par transduction de la souche β 5419 à l'aide d'un lysat du phage P1 récolté sur une souche auxiliaire (MS2131, Shapiro
 20 1990) portant un marqueur de résistance à la tétracycline dnaQ::miniTn10. Un tel clone résistant à la tétracycline et montrant un taux de mutation spontanée amplifié environ 1000 fois (pour l'acquisition de la résistance à la streptomycine) est cultivé à 30°C en milieu minimum glucose
 25 en présence de thymidine (0,3 mM). Après 24 h, les cellules sont récoltées, lavées deux fois dans un volume identique de milieu de culture sans thymidine. Un volume de 0,1 ml de la suspension résultante, correspondant à environ 10^8 bactéries, est étalé à la surface d'une série de boîtes de Pétri contenant une concentration de S-carbamyl-L-cystéine
 30 variant entre 0 et 8 mM par incrément de 1 mM additionnant du milieu minéral MS glucose (2 g/l). La même procédure est appliquée à la souche non-mutatrice β 5419, au gène dnaQ sauvage. L'ensemble des boîtes de Pétri est incubé 96 h à
 35 30°. Des colonies apparaissent sur les boîtes ayant une concentration de S-carbamyl-L-cystéine dépassant 2 mM dans

le seul cas où l'allèle mutateur dnaQ::miniTn10 a été introduit dans la souche testée.

Un lysat du phage P1 récolté à partir d'un tel clone est employé pour transduire la souche $\beta 5366$ portant l'allèle plasmidique thyA:Val146. Environ la moitié des transductants résistant à la kanamycine se montrent capables de croître en présence de 3 mM de S-carbamyl-L-cystéine et en absence de thymine ou de thymidine, parmi lesquels la souche $\beta 5455$. L'autre moitié des transductants en est incapable et exige la thymine ou la thymidine pour proliférer, tout comme la souche $\beta 5366$. Cette proportion entre les phénotypes s'accorde avec la distance génétique entre les locus vals et nrdD (0,4 mn). Ainsi, la suppression de la mutation faux-sens de thyA Cys->Val par une faible concentration de cystéine exogène pourrait résulter d'une altération du gène de la valyl-tRNA synthétase. Le locus vals d'une des souches obtenues par transduction de $\beta 5366$ et capables de croître en présence de 3 mM de S-carbamyl-L-cystéine et en absence de thymine ou de thymidine, désignée $\beta 5455$, est amplifié par réaction de polymérisation en chaîne et séquencé comme il est décrit dans l'Exemple 3. Un C changé en A cause le remplacement de la proline à la position 222 par la thréonine, confirmant ainsi la fixation d'une altération génétique dans le gène vals de la souche $\beta 5455$.

Exemple 6 : Sensibilité des mutants de la valyl-tRNA synthétase à des acides aminés non-canoniques.

Les souches $\beta 5455$, $\beta 8144$ et $\beta 8146$ sont testées pour leur sensibilité à des acides aminés artificiels qui présentent une ressemblance stérique avec la valine. Le test est réalisé sur des boîtes de milieu minéral MS glucose supplémenté avec de la thymidine. Les cellules sont cultivées en milieu aérobie (minéral MS glucose 0,3 mM

thymidine) pendant 24 h à 30°C et diluées au 1/250 dans du milieu minéral MS. 0,5 ml de cette suspension cellulaire sont étalés sur boîtes de Pétri contenant 25 ml de milieu minéral MS glucose. Un puits est ensuite évidé au centre de la boîte et rempli avec 0,1 ml d'une solution d'un acide aminé :

- (1) 100 mM L-2-amino-butyrate
- (2) 100 mM L-2-amino-valérate
- (3) 100 mM L-2-3-diamino-propionate
- 10 (4) 50 mM L-3-thiol-2-amino-butyrate.

Les boîtes sont ensuite incubées pendant 24 h à 30°C et l'apparition éventuelle sur les boîtes d'une zone d'inhibition autour du puits est enregistrée. Les diamètres des zones d'inhibition de croissance atténuées sur des boîtes de Pétri sont mesurés :

L-2-amino-butyrate : 5,2 cm (β 5455), 5,7 cm (β 8144), 6,7 cm (β 8146) ;

L-2-amino-valérate : 2,1 cm (β 5455), 1,5 cm (β 8144), 6,7 cm (β 8146) ;

20 L-2-3-diamino-propionate : 2,3 cm (β 5455), 2,7 cm (β 8144), 1,9 cm (β 8146) ;

L-3-thiol-2-amino-butyrate : 2,0 cm (β 5366), 4,6 cm (β 5455), 4,0 cm (β 8144), 4,0 cm (β 8146).

Le L-2-amino-butyrate, le L-2-amino-valérate et le L-2,3 diamino-propionate aux concentrations indiquées sont sans effet sur la souche β 5366 à l'allèle valS sauvage, mais inhibent la croissance des souches portant un gène valS muté. Le L-3-thiol-2-amino-butyrate inhibe la croissance de toutes les souches, mais on peut remarquer une inhibition plus importante sur les souches mutées. Ainsi, tout se passe comme si les mutants de la valyl-tRNA

synthétase avaient une spécificité élargie les rendant capables de charger des tRNA^{Val} avec des acides aminés qui ne peuvent être incorporés par la forme sauvage de l'enzyme.

BIBLIOGRAPHIE

- BAIN J.D., E.S. DIALA, C.G. GLABE, D.A. WACKER, M.H. LYTTLE, T.A. DIX and A.R. CHAMBERLIN, 1991 ; Site-specific
 5 incorporation of nonstructural residues during in vitro protein biosynthesis with semisynthetic aminoacyl-tRNAs, Biochemistry 30: 5411-5421.
- BOUZON, M. and P. MARLIERE, 1997 ; Human deoxycytidine kinase as a conditional mutator in Escherichia coli. C.R.
 10 Acad.Sci. Paris 320: 427-434.
- KUNKEL, T.A., and J.D. ROBERTS, 1987 ; Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. 154: 367-382.
- LEMEIGNAN, B., P. SONIGO and P. MARLIERE, 1993 ; Phenotypic
 15 suppression by incorporation of an alien amino acid. J. Mol. Biol. 231: 161-166.
- LEVH, T.F., J.C.TAYLOR and G.D. MARKHAM, 1988 ; The sulfate activation locus of Escherichia coli K12: cloning, genetic, and enzymatic characterisation. J. Biol. Chem. 263: 2409-
 20 2416.
- PARSOT, C., 1986 ; Evolution of biosynthetic pathways: a common ancestor for threonine synthase, threonine dehydratase and D-serine dehydratase. EMBO J., 5: 3013-3019.
- 25 RICHAUD, C., D. MENGIN-LECREULX, S. POCHET, E.J. JOHNSON, G.N. COHEN et al., 1993 ; Directed Evolution of Biosynthetic pathways. J. Biol. Chem. 268: 26827-26835.
- SAMBROOK, J., E.F. FRITSCH and T. MANIATIS, 1989 ; Molecular cloning: A Laboratory Manual. Cold Spring Harbor
 30 Laboratory Press, Cold Spring Harbor, NY.
- SANGER, F., S. NICKLEN and A.R. COULSON, 1977 ; DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74: 5463-5467.
- SHAPIRO, J.A 1990 ; Action of a transposable element in
 35 coding sequence fusions. Genetics 126: 293-299.

REVENDEICATIONS

1. Méthode permettant à des cellules d'acquérir la capacité de produire une protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel, caractérisée en ce qu'elle comprend les étapes suivantes :
- 5 a) la transformation desdites cellules par au moins une introduction d'une mutation faux-sens au niveau d'un codon cible d'un gène codant pour une protéine
- 10 nécessaire à la croissance desdites cellules, ladite protéine synthétisée à partir du gène ainsi muté n'étant plus fonctionnelle ;
- b) le cas échéant la culture des cellules obtenues à l'étape a) dans un milieu de culture contenant un
- 15 nutriment compensant la perte de fonctionnalité de ladite protéine ainsi mutée ; et
- c) la culture des cellules obtenues à l'étape a) ou b) dans un milieu de culture contenant l'acide aminé codé par ledit codon cible.
- 20
2. Méthode selon la revendication 1, caractérisée en ce que le milieu de culture de l'étape c) ne contient pas le nutriment nécessité par la perte de fonctionnalité de ladite protéine mutée.
- 25
3. Méthode selon l'une des revendications 1 et 2, caractérisée en ce que l'étape c) de culture desdites cellules comprend une série de cultures desdites cellules dans un milieu de culture contenant l'acide aminé codé par
- 30 ledit codon cible, chacune desdites cultures de la série étant effectuée jusqu'à obtention de la phase stationnaire de croissance et suivie d'un lavage des cellules obtenues,
-
- ~~le nombre de cultures de la série étant suffisant pour~~
- 35 permettre la sélection de mutations augmentant la suppression de ladite mutation faux-sens dudit gène muté.

4. Méthode selon l'une des revendications 1 à 3, caractérisée en ce que la mutation faux-sens est choisie parmi les mutations faux-sens qui ne réversent spontanément qu'à une très faible fréquence, de l'ordre d'un organisme
5 parmi au moins 10^{15} .
5. Méthode selon l'une des revendications 1 à 4, caractérisée en ce que la mutation faux-sens transforme un codon cible d'un gène codant pour une protéine nécessaire à
10 la croissance de ladite cellule en un codon qui comparativement au codon cible présente un changement d'au moins deux bases, de préférence trois bases.
6. Méthode selon l'une des revendications 1 à 5, caractérisée en ce que le codon cible code pour un acide
15 aminé de faible volume stérique.
7. Méthode selon l'une des revendications 1 à 6, caractérisée en ce que le codon cible code pour un acide
20 aminé amphiphile.
8. Méthode selon l'une des revendications 1 à 7, caractérisée en ce que le codon cible code pour un acide aminé dont volume stérique est inférieur ou sensiblement
25 égal au volume stérique de l'acide aminé codé par la mutation faux-sens.
9. Méthode selon l'une des revendications 5 à 8, caractérisée en ce que le codon cible code pour la
30 cystéine.
10. Méthode selon l'une des revendications 5 à 9, caractérisée en ce que l'acide aminé codé par la mutation faux-sens est la valine ou l'isoleucine.
- 35
11. Méthode selon l'une des revendications 1 à 10, caractérisée en ce que l'étape a) de transformation

desdites cellules est réalisée au moyen d'un vecteur comprenant une séquence dudit gène codant pour une protéine nécessaire à la croissance desdites cellules comportant ladite mutation faux-sens.

5

12. Méthode selon la revendication 11, caractérisée en ce que ledit vecteur est un vecteur plasmidique.

10

13. Méthode de sélection de cellules capables de produire une protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel caractérisée en ce qu'elle comprend les étapes a), le cas échéant b), et c) d'une méthode selon l'une des revendications 1 à 12, et la sélection des cellules capables de croître à l'étape c).

15

14. Méthode de sélection de cellules selon la revendication 13, caractérisée en ce qu'elle comprend en outre une étape d) de culture des cellules obtenues à l'étape c) dans un milieu de culture contenant ledit acide aminé codé par ledit codon cible, la concentration dudit acide aminé pouvant être à une concentration supérieure à la concentration dudit acide aminé utilisée à l'étape c), et le choix des cellules sensibles à la concentration dudit acide aminé utilisée à l'étape d).

25

15. Méthode de sélection de cellules selon l'une des revendications 13 et 14, caractérisée en ce que l'aminocyl-tRNA synthétase reconnaissant l'acide aminé codé par ladite mutation faux-sens desdites cellules sélectionnées est capable de charger sur un de ses tRNA associés un acide aminé non conventionnel ou un acide aminé autre que ledit acide aminé codé par ladite mutation faux-sens.

35

16. Méthode de sélection de cellules selon la revendication 15, caractérisée en ce que la séquence nucléique du gène codant pour ladite aminocyl-tRNA

synthétase comporte au moins une mutation comparée à la séquence du gène sauvage correspondant.

5 17. Méthode de sélection de cellules selon la revendication 16, caractérisée en ce que ladite mutation n'a pas été introduite par une technique de recombinaison génétique.

10 18. Cellule obtenue par une méthode selon l'une des revendications 1 à 17.

15 19. Cellule isolée capable de produire une protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel, caractérisée en ce qu'elle comprend une aminoacyl-tRNA synthétase reconnaissant un acide aminé donné capable de charger sur un de ses tRNA associés un acide aminé non conventionnel ou un acide aminé autre que ledit acide aminé donné, et en ce que la séquence nucléique du gène codant pour ladite aminoacyl-tRNA synthétase comporte au moins une mutation comparée à la séquence du gène sauvage correspondant, ladite mutation n'ayant pas été introduite par une technique de recombinaison génétique.

25 20. Cellule selon la revendication 18 ou 19, caractérisée en ce qu'il s'agit d'une cellule procaryote ou eucaryote.

21. Cellule selon la revendication 20, caractérisée en ce qu'il s'agit d'une cellule procaryote.

30

22. Cellule selon l'une des revendications 18 à 21, caractérisée en ce qu'elle est choisie parmi les cellules suivantes déposées à la CNCM (Collection Nationale de Culture de Microorganismes, Paris, France) :

35 a) souche *E. coli* déposée à la CNCM sous le N° I-2025 le 25 mai 1998,

- b) souche *E. coli* déposée à la CNCM sous le N° I-2026 le 25 mai 1998, et
c) souche *E. coli* déposée à la CNCM sous le N° I-2027 le 25 mai 1998.

5

23. Utilisation d'une méthode selon l'une des revendications 1 à 17 pour la production de protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel.

10

24. Utilisation d'une cellule selon l'une des revendications 18 à 22 pour la production de protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel.

15

25. Procédé de production d'une protéine dont la séquence d'acides aminés comprend au moins un acide aminé non conventionnel, caractérisé en ce qu'il comprend les étapes suivantes :

- 20 a) le cas échéant, la sélection d'une cellule par une méthode selon l'une des revendications 13 à 17 ;
b) la culture de ladite cellule sélectionnée à l'étape a) ou d'une cellule selon l'une des revendications 18 à 22 dans un milieu de culture et des conditions de culture
25 permettant la croissance de ladite cellule ; et
c) l'isolement de ladite protéine comprenant au moins un acide aminé non conventionnel à partir du surnageant de culture et/ou du culot cellulaire obtenu à l'étape b).

30 26. Procédé selon la revendication 25, caractérisé en ce que ledit milieu de culture de l'étape b) permettant la croissance de ladite cellule contient ledit acide aminé non conventionnel ou un de ses précurseurs.

35 27. Procédé selon la revendication 25, caractérisé en ce que ledit acide aminé non conventionnel est synthétisé par ladite cellule.

28. Procédé selon la revendication 27, caractérisé en ce que la synthèse dudit acide aminé non conventionnel est augmentée par modification génétique de ladite cellule.

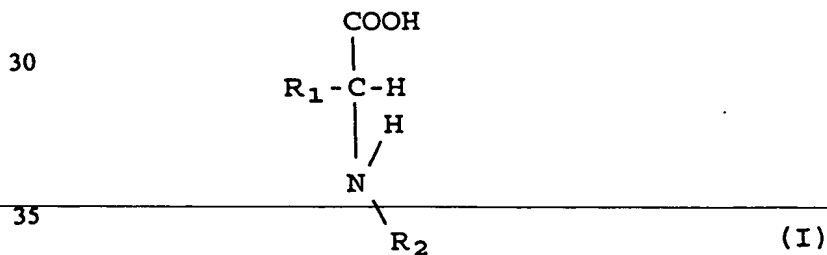
5 29. Procédé selon l'une des revendications 25 à 28, caractérisé en ce que ladite cellule est auxotrophe pour l'acide aminé codé par ledit codon cible.

10 30. Procédé selon l'une des revendications 25 à 29, caractérisé en ce que ladite cellule comprend un gène d'intérêt homologue ou hétérologue dont la séquence codante comporte au moins un codon cible.

15 31. Procédé selon la revendication 30, caractérisé en ce que l'étape b) comprend les composés nécessaires à l'induction de la synthèse de la protéine codée par ledit gène d'intérêt.

20 32. Procédé selon la revendication 30 ou 31, caractérisé en ce que l'activité biologique de la protéine codée par ledit gène d'intérêt est au moins partiellement conservée après l'incorporation dudit acide aminé non conventionnel au niveau du codon cible dudit gène d'intérêt.

25 33. Procédé selon l'une des revendications 25 à 32, caractérisé en ce que l'acide aminé non conventionnel est choisi parmi les acides aminés non conventionnels de formule I de configuration L



dans laquelle :

R₁ ou R₂ représente des radicaux contenant un groupe fonctionnel capable de réagir de manière sélective.

34. Procédé selon la revendication 33, caractérisé en ce que le groupe fonctionnel est choisi parmi les groupes aldéhyde, cétone, éthényle, éthyne ou nitrile.

5 35. Procédé selon l'une des revendications 25 à 34, pour la fonctionnalisation de protéine.

36. Procédé de purification de protéine, caractérisé en ce qu'il comprend les étapes suivantes :

- 10 a) l'incorporation dans la séquence d'acides aminés de ladite protéine d'un acide aminé non conventionnel contenant un groupe fonctionnel capable de réagir de manière sélective par un procédé selon l'une des revendications 25 à 35 ;
- 15 b) la mise en contact de la solution contenant la protéine obtenue à l'étape a) avec un support comprenant un composé capable de réagir spécifiquement avec ledit groupe fonctionnel et de fixer spécifiquement ladite protéine ; et
- 20 c) l'isolement de ladite protéine fixée sur le support.

37. Procédé de fixation d'une protéine sur un composé chimique ou biochimique, caractérisé en ce qu'il comprend les étapes suivantes :

- 25 a) l'incorporation dans la séquence d'acides aminés de ladite protéine par un procédé selon l'une des revendications 25 à 35 d'un acide aminé non conventionnel contenant un groupe fonctionnel capable de réagir de manière sélective ;
- 30 b) la mise en contact de la protéine obtenue à l'étape a) avec ledit composé chimique ou biochimique comprenant un groupe capable de réagir spécifiquement avec ledit groupe fonctionnel dans un milieu permettant la réaction.

35

38. Procédé selon la revendication 37, caractérisé en ce que ledit composé chimique ou biochimique est lui-même fixé

sur un support solide ou est un composé constitutif d'un support solide.

39. Procédé selon la revendication 37 pour la préparation
5 d'un complexe protéique.

40. Procédé selon la revendication 39, caractérisé en ce
que la protéine fixée ou le composé chimique ou biochimique
est choisi parmi les composés thérapeutiques, cosmétiques
10 ou diagnostiques.

41. Procédé selon la revendication 39 ou 40, caractérisé
en ce que le composé chimique ou biochimique est choisi
parmi les composés capables de modifier l'activité
15 biologique de la protéine fixée.

42. Procédé selon la revendication 39 ou 40, caractérisé
en ce que le composé chimique ou biochimique est choisi
parmi les composés dont l'activité biologique peut être
20 modifiée par la protéine fixée.

43. Procédé selon l'une des revendications 39 à 42,
caractérisé en ce que le composé chimique ou biochimique
est choisi parmi les composés comprenant une protéine, un
25 polynucléotide, un acide gras, un sucre ou un polymère
naturel ou synthétique.

44. Protéine obtenue par un procédé selon l'une des
revendications 25 à 36.
30

45. Protéine selon la revendication 44, caractérisée en
ce qu'il s'agit d'une protéine recombinante.

46. Complexe protéique obtenu par un procédé selon l'une
35 revendications 39 à 43.

47. Utilisation d'une protéine selon la revendication 44 ou 45, ou d'un complexe protéique selon la revendication 46 comme réactif de diagnostic.

5 48. Procédé de diagnostic, caractérisé en ce qu'il met en oeuvre une protéine selon la revendication 44 ou 45, ou un complexe protéique selon la revendication 46.

49. Nécessaire de diagnostic, caractérisé en ce qu'il
10 contient une protéine selon la revendication 44 ou 45, ou un complexe protéique selon la revendication 46.

50. Utilisation d'une protéine selon la revendication 44 ou 45, d'un complexe protéique selon la revendication 46 ou
15 d'une cellule selon l'une des revendications 18 à 22 pour la préparation d'une composition pharmaceutique ou cosmétique.

51. Composition pharmaceutique ou cosmétique comprenant
20 une protéine selon la revendication 44 ou 45, un complexe protéique selon la revendication 46 ou une cellule selon l'une des revendications 18 à 22.

Selon l'invention, l'étape c) de culture desdites cellules peut comprendre une série de cultures desdites cellules dans un milieu de culture contenant l'acide aminé codé par ledit codon cible, chacune desdites cultures de la
5 série étant effectuée jusqu'à obtention de la phase stationnaire de croissance et suivie d'un lavage des cellules obtenues, le nombre de cultures de la série étant suffisant pour permettre la sélection de mutations augmentant la suppression de ladite mutation faux-sens dudit gène muté et
10 la propagation de l'allèle correspondant audit gène muté.

L'invention concerne en outre une méthode selon l'invention, caractérisée en ce que la mutation faux-sens est choisie parmi les mutations faux-sens qui ne réversent spontanément qu'à une très faible fréquence, de l'ordre d'un
15 organisme parmi au moins 10^{15} .

De préférence, la mutation faux-sens sera choisie parmi les mutations faux-sens qui transforment un codon cible d'un gène codant pour une protéine nécessaire à la croissance de ladite cellule en un codon qui comparative-ment au codon
20 cible présente un changement d'au moins deux bases, de manière plus préférée trois bases.

On préfère également les méthodes selon l'invention, caractérisées en ce que le codon cible code pour un acide aminé de faible volume stérique et/ou amphiphile et/ou de
25 volume stérique inférieur ou sensiblement égal au volume stérique de l'acide aminé codé par la mutation faux-sens.

Parmi les codons cibles, on préfère notamment les codons cibles codant pour la cystéine et les mutations faux-sens choisies parmi les mutations faux-sens qui transforment un
30 codon cible en un codon codant pour la valine ou l'isoleucine.

L'invention concerne en outre une méthode selon l'invention, caractérisée en ce que l'étape a) de ~~transformation desdites cellules est réalisée au moyen d'un~~
35 vecteur comprenant une séquence dudit gène codant pour une protéine nécessaire à la croissance desdites cellules comportant ladite mutation faux-sens, notamment au moyen d'un vecteur plasmidique.